A Guide to Good Data Management

How to build an intelligent data management strategy

It all starts with data. In an era where data-driven and analytics-enabled decisions fuel business growth and innovation, organizations realize data is the key to a competitive edge. But to unlock the full potential of data, you have to manage it strategically.

An effective data management strategy enables you to find the right data, at the right time and in the right place—so you can quickly extract the most valuable insights. Chances are, however, that the methodology you use today will not achieve these outcomes.

Here's why:

- The "IT as usual," centralized data management approach no longer meets the realities of the data world.
- Consumers expect real-time access and interaction, and your organization's expectations have changed accordingly.
- Most approaches to data management prevent you from creating value out of the growing quantity of fragmented data you're collecting and processing.

An **effective data management strategy** enables you to find the right data, at the right time and in the right place

To go beyond the barriers to intelligent data management, you need to understand why these hurdles occur. This white paper takes a look at what's preventing your organization from creating value out of data, what you need to consider in your data management strategy, and what questions you should ask your teams or potential partners.

93% of businesses face **entrenched barriers to transformation**, and identified the inability to extract valuable insights from data/information overload as the top third barrier.[2]

Gartner forecasts that by 2022, **more than 50% of enterprise data will be created and processed outside the data center or cloud**, up from less than 10% in 2019.[3]

The problem with today's data management

Decision makers understand the importance of data, but there's a gap in their ability to achieve desired outcomes.

89% of organizations say data management and analytics is one of their **top ten business and IT priorities** for the next 24 months.[1]

The inability to view data as a single source of truth stems from data management strategies that are optimized for a centralized IT infrastructure.

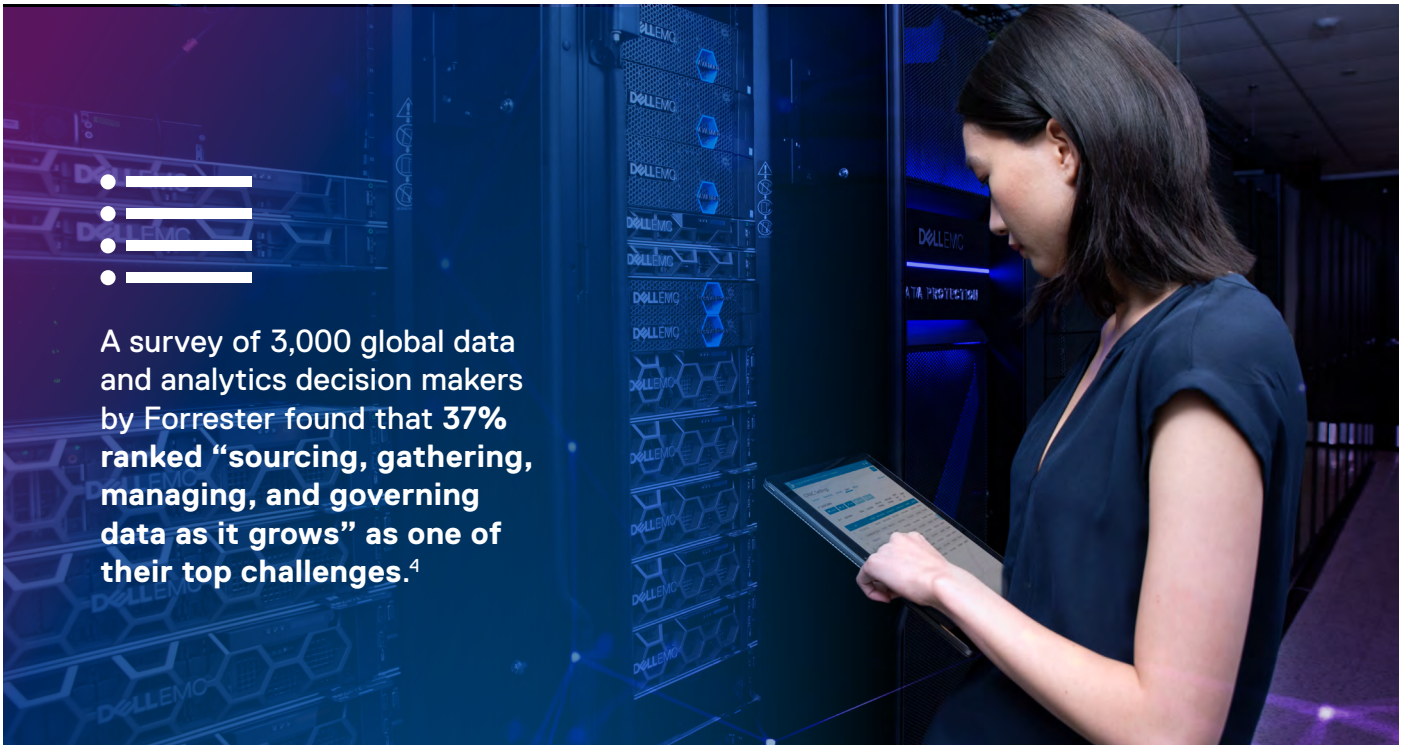
Traditionally, data is transferred to a central location, aggregated in data warehouses and data lakes and eventually processed in batches. But organizations are digitizing more business processes and activities—and they increasingly require real-time data analytics and actionable insights.

Transferring large quantities of data in real time to a central server for processing is cumbersome and expensive. There's simply too much data to continuously send to the cloud or data center. Additionally, problems due to connectivity, latency and bandwidth negatively impact operational performance.

The majority of data growth is now taking place at the edge. As data becomes more distributed, it will further underscore the limitations of existing data management models and the advantages of edge computing.

To help you realize the full potential of data, your data management strategy needs to consider the following three challenges:

- Data—unlocking value from new data types and realities.
- Infrastructure—achieving an infrastructure and data management strategy that are consistent from the core to the cloud to the edge.
- People—implementing a coordinated data management approach with a common goal.



Data: Embracing rather than eliminating data silos

Data volume and velocity are surging. Organizations manage 13.53 petabytes of data on average, a staggering 831% increase since 2016.[5] This skyrocketing growth rate exceeds organizations' ability to harness the value of their data.

With data becoming more distributed, many organizations attempt to eliminate silos by using the conventional approach of transferring data to the cloud or data center. However, the centralized infrastructure can't keep up with the growing number of edge locations and the data volumes they generate. Nor can it accommodate those use cases that require real-time insights for quick action. Embracing data silos rather than eliminating them is the only way to move forward.

Your strategy needs to ensure that:

- All your data—from different sources and sites—can be collected, stored and archived regardless of whether it resides on-premises, is stored in multiple clouds or is highly distributed across a myriad of edge locations.
- Your data engineers, data scientists and data analysts can access the data easily and efficiently to deliver high-quality, production-ready data sets, and then develop machine-learning algorithms and analyze the data to derive meaningful insights.
- As the attack surface expands due to the growing amount of data and locations, you can maintain regulatory compliance and security.

Key questions to ask:

- **How are you cataloging data?**
  Are you able to easily identify, classify and move data between heterogeneous storage systems and the cloud?
- **Can you discover unknown, dormant data?**
  Most of the data that organizations collect, store and archive isn't being used to influence business decisions. Do you have the ability to locate and correlate data from all the different sources and silos, as well as to grant access to the right people? How easily and quickly can you do so?
- **Are you meeting regulatory and compliance requirements?**
  Are you using advanced data protection, security and governance controls to help you comply with new regulations and reduce your risks? How confident are you that you're aligned to the latest mandates?

The infrastructure evolution: Optimizing your IT for data-centric workloads

In response to the growing data volumes and the need for fast network speed and reliable uptime, data-driven organizations are moving their infrastructure closer to the edge. By processing, analyzing and acting on the data closer to the locations that generate it, you don't just solve the inefficiency of streaming the data to the cloud or data center. You can establish enhanced security protocols, address compliance issues and lower costs.

But this move poses a new challenge: How can you achieve a consistent infrastructure and data management across your entire IT landscape: on-premises, across multiple clouds and at the edge?

Your strategy needs to ensure that:

- The infrastructure can respond in an agile way to data processing requirements.
- You're taking advantage of the benefits of a distributed approach without the drawbacks of siloed data management.
- Your technology landscape is optimized for machine learning workloads, for faster time to data insights.

Key questions to ask:

- **How are you integrating internal and external data sources across all locations?**
  Are you managing data consistently, whether the data sources are located on-premises, in the cloud or at the edge?
- **How well does your infrastructure support your storage needs?**
  Does your infrastructure meet the demands for high-capacity, high-performance and highly resilient storage?
- **Is your infrastructure optimized for data-intensive workloads?**
  Do you have sufficient, affordable processing power for sophisticated, advanced techniques such as artificial intelligence and machine learning?

Only **22%** of data management teams' time is spent on innovation.[6]

The people challenge: Coordinating data roles and skill sets

The individuals and teams involved in advancing your organization through the data lifecycle represent different business units and departments. They bring diverse skills to the table and their own set of objectives. This may prevent you from having a coordinated data management approach with common goals.

The competing goals and lack of collaboration create inefficiencies in your data pipeline. Yet, many organizations are missing key data roles and skill sets, and the shortage of highly skilled and highly specialized talent further compounds the problem.

Data scientists are a rare and expensive resource for your organization. You can't afford for them to be distracted by having to create data pipelines or scaling infrastructure. And even when they're focused on their own responsibilities, are they spending too much time away from core activities because they're trying to identify relevant data sources and gain access to the data?

Your strategy needs to ensure that:

- ✓ You have the right people and expertise to architect and implement artificial intelligence (AI) solutions that support your business objectives.
- ✓ You're streamlining processes and coordinating goals so data scientists can spend more time on innovation and monetization rather than improving the quality of data.

Key questions to ask:

- ✓ **How are you streamlining data ingestion and application development?** *Are you getting business-ready data when it's timely, so you can act faster on insights? What portion of your team's time is spent on cleansing, aggregating or preparing data for analysis, rather than creating value with it?*
- ✓ **Do you have the people and expertise to determine use cases?** *Can you align business and IT objectives to AI to ensure you have the right use cases to develop?*

What's next

Competing and growing in the data era requires you to become an intelligent business. You can't afford to maintain the status quo if you want to move your organization forward.

Your modern data management strategy needs to consider all aspects—your data, infrastructure and people—for successful outcomes and continued relevance to customers. But you don't need to reinvent the wheel. Read 'Creating an intelligent business' for a tested, proven data management blueprint that you can take to a potential partner.

Click here to learn how data management can create a more intelligent business.

Click here to read why you should be rethinking data management.

DELL Technologies

Intel® Innovation Built-in

intel®

DellTechnologies.com/ConnectedCIO

Sources:

1. ESG eBook "The Data Management Imperative: The Critical Differentiator in the Data Era" commissioned by Dell Technologies, October 2020. Results based survey of 500 business leaders in North America and Western Europe on data management maturity and practices. Actual results will vary. Full report: https://www.delltechnologies.com/en-us/what-we-do/connectedcio.htm#overlay=/collaterals/unauth/brochures/solutions/the-data-management-imperative.pdf
2. Dell Technologies Digital Transformation Index, October 2020. Research conducted by Vanson Bourne surveying 4,300 business leaders across 18 countries. https://www.delltechnologies.com/dtindex
3. Smarter with Gartner, "Gartner Top 10 Trends Impacting Infrastructure & Operations for 2020," 10 December 2019. https://www.gartner.com/smarterwithgartner/gartner-top-10-trends-impacting-infrastructure-operations-for-2020/.
4. Forrester Analytics Global Business Technographics Data And Analytics Survey, June 2019. https://www.forrester.com/Global+Business+Technographics+Data+And+Analytics+Survey+2019/-/E-sus5171
5. Vanson Bourne, Global Data Protection Index 2020 Snapshot, March 2020. A commissioned survey of 1,000 IT decision makers across 15 countries globally on their data protection strategies, their approaches to data protection in cloud environments, and the relative preparedness of their businesses in cases of disruption. https://www.delltechnologies.com/en-us/data-protection/gdpi/index.htm#gdpi_2020
6. Gartner, Survey Analysis: Data Management Struggles to Balance Innovation and Control, Melody Chien, Nick Heudecker, 19 March 2020. https://www.gartner.com/en/documents/3982237/survey-analysis-data-management-struggles-to-balance-inn